

## SEQUENCE LISTING

<110> MCGILL UNIVERSITY  
 SZYF, Moshe  
 BHATTACHARYA, Sanjoy K.  
 RAMCHANDANI, Shyam

<120> DNA DEMETHYLASE, THERAPEUTIC AND  
 DIAGNOSTIC USES THEREOF

<130> 1770-183 "PCT" FC/ld

*Sub B*

<150> CA 2,220,805  
<151> 1997-11-12

<150> CA 2,230,991  
<151> 1998-05-11

<160> 10

<170> FastSEQ for Windows Version 3.0

<210> 1  
<211> 1804  
<212> DNA  
<213> Unknown

<400> 1

ccgcctctgct	ggcgaaaaaa	gtctccggaa	ttccaaaggc	tccgttacgg	aagaagcgca	60
gagccggctg	gggggggggg	tggatgcgg	cgcaccgggg	gggaggccgc	tgcgtccccgg	120
agcaggagga	ggggggagagc	gcggcgaaaa	gcagcggcgc	tggcgccgac	tccgcctatag	180
agcagggggg	ccagggcagc	gcgcgtcgct	cgtccccgt	gagcggcggt	cgcaggaaag	240
gcccgtcgaa	cgccggccgt	ggccgggggg	ggtggaaagca	ggcgccccgg	ggcgccggcg	300
tctgtggccg	tggccgtggc	cgtggccggg	gtcgccggcg	tggccggggc	cgggggccggg	360
gccgcggccg	tcccccagat	ggccggcagcg	gccttggcg	cgacggcgcc	ggcgccgcgg	420
gcggctgcgg	cgtcgccagc	ggtggcgccg	tgcggccccc	gccccatcct	gtccctttcc	480
cgtcgccgg	ctcgccggcc	ggggcccgagg	gaccccgggc	cacggagagc	ggaaagagga	540
tggactgccc	ggccctcccc	cccgatgg	agaaggagga	agtgtatccga	aatcaggc	600
tcagtgtcg	caagagcgat	gtctactact	tcaatccaag	tggtaagaag	ttcagaagta	660
aacctcagct	ggcaagatac	ctggggaaatg	ctgttgacct	tagcagtttt	gacttcagga	720
ccggcaagat	gatgcctagt	aaattacaga	agaacaagca	gagactccgg	aatgaccccc	780
tcaatcagaa	caagggtaaa	ccagaccta	acacaacatt	gccaattaga	caaactgcat	840
caatttcaa	gcaaccagta	accaaaattca	cgaaccaccc	gagaataag	gtgaagttag	900
accccccacg	gatgaatgaa	caaccacgtc	agcttttctg	ggagaagagg	ctacaaggac	960
ttagcgcata	agatgtaca	gaacaaatta	aaaaaacat	ggagctacct	aaaggcttc	1020
aaggagtcgg	tccaggttag	aatgacgaga	cccttctgtc	tgcgtggcc	agtgcattac	1080
acacaagctc	tgcgcctatc	acaggacaag	tctctgtgc	cgtggaaaag	aaccctgctg	1140
tttggcttaa	cacatctaa	ccccctctgca	aagctttcat	tgttacagat	gaagacatta	1200
ggaaaacagga	agagcgagtc	caacaagtac	gcaagaaaact	ggaggaggca	ctgatggccg	1260
acatcctgtc	ccgggctgct	gacacggagg	aagttagacat	tgacatggac	agtggagatg	1320
aggcgtaaga	atatgatcag	gtaactttcg	actgacccttc	cccaagagca	aattgttaga	1380
aacagaattt	aaacatttcc	actgggtttc	gcctgttaaa	aaaagtgtac	ctgagcacat	1440
agcttttaa	tagcaactaac	caatgcctt	ttagatgtat	tttgatgtat	tatatctatt	1500
attccaaatg	atgtttat	tgaatcctag	gactaaaaat	gagtctttta	taatagcaag	1560
caggccctt	ccggcgtcagt	gcagcttga	ggccaggtgc	agtctactgg	aaaggtagca	1620
cttacgtgaa	atatttgg	ccccacagt	tttaatataa	acagatcagg	agtaccaaataat	1680

aagttccca attaaagatt attatacttc actgtatata aacagatttt tatactttat 1740  
tgaagaaga tacctgtaca ttcttccatc atcactgtaa agacaaataa atgactatat 1800  
tcac 1804

<210> 2  
<211> 411  
<212> PRT  
<213> Unknown

*JKB*

<400> 2  
Met Arg Ala His Pro Gly Gly Gly Arg Cys Cys Pro Glu Gln Glu Glu  
1 5 10 15  
Gly Glu Ser Ala Ala Gly Gly Ser Gly Ala Gly Gly Asp Ser Ala Ile  
20 25 30  
Glu Gln Gly Gly Gln Gly Ser Ala Leu Ala Pro Ser Pro Val Ser Gly  
35 40 45  
Val Arg Arg Glu Gly Ala Arg Gly Gly Arg Gly Arg Gly Arg Trp  
50 55 60  
Lys Gln Ala Gly Arg Gly Gly Val Cys Gly Arg Gly Arg Gly Arg  
65 70 75 80  
Gly Arg  
85 90 95  
Pro Pro Ser Gly Gly Ser Gly Leu Gly Gly Asp Gly Gly Cys Gly  
100 105 110  
Gly Gly Gly Ser Gly Gly Gly Ala Pro Arg Arg Glu Pro Val Pro  
115 120 125  
Phe Pro Ser Gly Ser Ala Gly Pro Gly Pro Arg Gly Pro Arg Ala Thr  
130 135 140  
Glu Ser Gly Lys Arg Met Asp Cys Pro Ala Leu Pro Pro Gly Trp Lys  
145 150 155 160  
Lys Glu Glu Val Ile Arg Lys Ser Gly Leu Ser Ala Gly Lys Ser Asp  
165 170 175  
Val Tyr Tyr Phe Ser Pro Ser Gly Lys Lys Phe Arg Ser Lys Pro Gln  
180 185 190  
Leu Ala Arg Tyr Leu Gly Asn Thr Val Asp Leu Ser Ser Phe Asp Phe  
195 200 205  
Arg Thr Gly Lys Met Met Pro Ser Lys Leu Gln Lys Asn Lys Gln Arg  
210 215 220  
Leu Arg Asn Asp Pro Leu Asn Gln Asn Lys Gly Lys Pro Asp Leu Asn  
225 230 235 240  
Thr Thr Leu Pro Ile Arg Gln Thr Ala Ser Ile Phe Lys Gln Pro Val  
245 250 255  
Thr Lys Val Thr Asn His Pro Ser Asn Lys Val Lys Ser Asp Pro Gln  
260 265 270  
Arg Met Asn Glu Gln Pro Arg Gln Leu Phe Trp Glu Lys Arg Leu Gln  
275 280 285  
Gly Leu Ser Ala Ser Asp Val Thr Glu Gln Ile Ile Lys Thr Met Glu  
290 295 300  
Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Gly Ser Asn Asp Glu Thr  
305 310 315 320  
Leu Leu Ser Ala Val Ala Ser Ala Leu His Thr Ser Ser Ala Pro Ile  
325 330 335  
Thr Gly Gln Val Ser Ala Ala Val Glu Lys Asn Pro Ala Val Trp Leu  
340 345 350  
Asn Thr Ser Gln Pro Leu Cys Lys Ala Phe Ile Val Thr Asp Glu Asp  
355 360 365

Ile	Arg	Lys	Gln	Glu	Glu	Arg	Val	Gln	Gln	Val	Arg	Lys	Lys	Leu	Glu
370					375						380				
Glu	Ala	Leu	Met	Ala	Asp	Ile	Leu	Ser	Arg	Ala	Ala	Asp	Thr	Glu	Glu
385					390				395					400	
Met	Asp	Ile	Glu	Met	Asp	Ser	Gly	Asp	Glu	Ala					
					405				410						

<210> 3

<211> 1589

<212> DNA

<213> Unknown

<400> 3

cacgcgcggg	cgggtgggccc	gagcggcccc	cctagcgggg	gctgtgaagc	gcgggggaggc	60
ggccgagcgg	gtggcgaagc	cgcgccgcgc	ccggctgggg	gcggagggcg	gaggcccgtg	120
ggacagaaca	gctgcggcga	gtggcggcgg	cgagggagc	cgaatcggcg	acgagcccg	180
gggtcgaac	ttgcagaagd	ggcgccggcg	cgccatcg	ccacggcggg	cgaaaaagcc	240
ggggcgaat	ggagcggaaag	agggtggagt	gcccgccgt	cccgcagggc	tggaaaaggg	300
aagaagtgc	caggaggtcg	gggctgtcgg	ccggccacag	ggatgtctt	tactatagcc	360
ccagcggaa	gaagttccgc	agcaagccac	aactggcacg	ttacctggc	ggatccatgg	420
acctcagcac	cttcgacttc	cgcacccgaa	agatgttgat	gaacaagatg	aataagatc	480
gccagcgtgt	gcgcstatgt	tcttccaacc	aggtaaggg	caagcctgac	ctgaacaccg	540
cgctgcctgt	acggcagact	gcatccatct	tcaagcaacc	ggtgaccaag	atcaccacc	600
accccaagcaa	caaggtcaag	agcggccgc	agaaggcagt	ggaccagccg	aggcagctt	660
tctgggagaa	gaagctaagt	ggatggagt	ccttgacat	tgcagaagaa	ctggtagg	720
ccatggactt	gcccaagggc	ctgcaggag	tggccctgg	ctgtacagat	gagacgtgc	780
tgtcagccat	tgcgagtgct	ctacacacca	gcaccctg	cattacaggc	cagctctg	840
cagccgtgga	gaagaaccct	ggtgtgtggc	tgaacactgc	acagccactg	tgcaaagct	900
tcatggtgac	agatgacgac	atcaggaagc	aggaggagct	ggtacagcag	gtacggaaagc	960
gcctggagga	ggcactgtatg	gccgacatgc	tagctcatgt	ggaggagctt	gcccggagacg	1020
gggaggcacc	actggacaag	gcctgtgcag	aggaggaaga	ggaggaggaa	gaggaggagg	1080
aagagccgga	gcccagagcga	gtgttagcaca	ggtgcctcgt	ccaagtctgg	gctgcagact	1140
gccttcagcc	ttgcctggac	caggttagggg	ccagacctgt	aggaggcagc	cgtccacctc	1200
cttccaaag	cctcctgtct	ccaggtctca	gtgcaggag	ccctgtgga	ccttgaactc	1260
acttgtccct	gctgtgcctg	gcaggaagcc	ccacactgaa	agcagatgag	cagtgaccca	1320
actgagaggc	cacccggaca	cagtcacctc	cctgcctcct	tatcatagga	caaggccttg	1380
cttggcaccg	aggagctggg	agccgtgtt	ggtgctggag	gaagtttctg	gaaacacacc	1440
tggctatgcc	caccttatgt	ccctaaggct	attacaggcc	aggtttgg	ctgctccggc	1500
ccacagggct	gcccagcctc	cccacactga	ggtgcagcag	cccaccagga	agtcaacttc	1560
cttcaataaa	ctgtatggtag	gaacttgt				1589

<210> 4

<211> 291

<212> PRT

<213> Unknown

<400> 4

```

Met Glu Arg Lys Arg Trp Glu Cys Pro Ala Leu Pro Gln Gly Trp Glu
      1          5          10          15
Arg Glu Glu Val Pro Arg Arg Ser Gly Leu Ser Ala Gly His Arg Asp
      20         25         30
Val Phe Tyr Tyr Ser Pro Ser Gly Lys Lys Phe Arg Ser Lys Pro Gln
      35         40         45
Leu Ala Arg Tyr Leu Gly Gly Ser Met Asp Leu Ser Thr Phe Asp Phe
      50         55         60
Arg Thr Gly Lys Met Leu Met Ser Lys Met Asn Lys Ser Arg Gln Arg
      65         70         75         80

```

Val Arg Tyr Asp Ser Ser Asn Gln Val Lys Gly Lys Pro Asp Leu Asn  
 85 90 95  
 Thr Ala Leu Pro Val Arg Gln Thr Ala Ser Ile Phe Lys Gln Pro Val  
 100 105 110  
 Thr Lys Ile Thr Asn His Pro Ser Asn Lys Val Lys Ser Asp Pro Gln  
 115 120 125  
 Lys Ala Val Asp Gln Pro Arg Gln Leu Phe Trp Glu Lys Lys Leu Ser  
 130 135 140  
 Gly Leu Asn Ala Phe Asp Ile Ala Glu Glu Leu Val Lys Thr Met Asp  
 145 150 155 160  
 Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Gly Cys Thr Asp Glu Thr  
 165 170 175  
 Leu Leu Ser Ala Ile Ala Ser Ala Leu His Thr Ser Thr Met Pro Ile  
 180 185 190  
 Thr Gly Gln Leu Ser Ala Ala Val Glu Lys Asn Pro Gly Val Trp Leu  
 195 200 205  
 Asn Thr Thr Gln Pro Leu Cys Lys Ala Phe Met Val Thr Asp Glu Asp  
 210 215 220  
 Ile Arg Lys Gln Glu Glu Leu Val Gln Gln Val Arg Lys Arg Leu Glu  
 225 230 235 240  
 Glu Ala Leu Met Ala Asp Met Leu Ala His Val Glu Glu Leu Ala Arg  
 245 250 255  
 Asp Gly Glu Ala Pro Leu Asp Lys Ala Cys Ala Glu Asp Asp Asp Glu  
 260 265 270  
 Glu Asp Glu Glu Glu Glu Glu Pro Asp Pro Asp Pro Glu Met  
 275 280 285  
 Glu His Val  
 290

<210> 5  
 <211> 1966  
 <212> DNA  
 <213> Unknown

<400> 5

gggggcgtgg ccccgagaag gcggagacaa gatggccgccc catagcgctt ggaggaccta 60  
 agaggcggtg gccggggcca cgccccgggc aggaggggccg ctctgtgcgc gcccgcctca 120  
 tcatgtttgc gcgcgtcccc cgcgccgcgc gtcgcggcg gggcggtct ccgggattcc 180  
 aagggctcggt ttacggaaga agcgcagcgc cggctggga gggggctgga tgcgcgcgca 240  
 cccggggggaa ggccgctgct gcccggagca ggaggagggg gagagtgcgg cggggggcag 300  
 cggcgctggc ggcgactccg ccatagagca gggggggccag ggcagcgcgc tcgccccgtc 360  
 cccgggtgagc ggcgtgcgca gggaaaggcgc tcggggccgc gggcgtggcc gggggcggtg 420  
 gaagcaggcg ggcggggcg gccgcgtctg tggccgtggc cggggccggg gccgtggccg 480  
 gggacgggaa cggggccggg gccggggccg cggccgtccc ccgagtggcg gcagcggcct 540  
 tggccggcgc ggcggcggt gccgcggcg cggcagcggt ggcggcggcg cccccccggc 600  
 ggagccggtc cttttccgt cggggagcgc gggggccggg cccaggggac cccggggcac 660  
 ggagagcggg aagaggatgg attgcggc cttttttttt ggatgaaaga aggaggaagt 720  
 gatccgaaaa tctgggctaa gtgtggcaa gagcgatgtc tactacttca gtccaagtgg 780  
 taagaagttc agaagcaagc ctcagttggc aaggtacctg ggaaatactg ttgatctcag 840  
 cagttttgac ttccagaactg gaaagatgtat gccttagaaa ttacagaaga acaaacagag 900  
 actgcgaaac gatcctctca atcaaaataa gggtaaaacca gacttgaata caacattgcc 960  
 aatttagacaa acagcatcaa ttttcaaaca accggtaacc aaagtcaaa atcatcctag 1020  
 taataaaatgt aaatcagacc cacaacgaat gaatgaacag ccacgtcagc ttttctggaa 1080  
 gaagaggctaa caaggactta gtgcattcaga tgtaacagaa caaattataa aaaccatgg 1140  
 actacccaaa ggtcttcaag gagttggtcc aggttagcaat gatgagaccc ttttatctgc 1200  
 tggccgttgcacttgc gccaatcaca gggcaagtct ccgctgtgt 1260  
 gggaaaagaac cctgctgttt ggcttaacac atctcaaccc ctctgcaaaag cttttattgt 1320

cacagatgaa gacatcagga aacaggaaga gcgagtgacag caagtacgca agaaattgga 1380  
 agaagcactg atggcagaca tcttgcgcg agctgctgat acagaagaga tggatattga 1440  
 aatggacagt ggagatgaag cctaagaata tgatcaggta actttcgacc gactttcccc 1500  
 aagrgaaaaat tcctagaaaat tgaacaaaaa tgtttccact ggctttgcc tgtaagaaaa 1560  
 aaaatgtacc cgagcacata gagctttta atagcactaa ccaatgcctt ttttagatgta 1620  
 ttttgcgtt atatatctat tattcaaaaa atcatgttta tttagtcc taggacttaa 1680  
 aattagtctt ttgtaatatac aaggcaggacc ctaagatgaa gctgagctt tgatgccagg 1740  
 tgcaatctac tggaaatgta gcacttacgt aaaacatttgc tttcccccac agtttaata 1800  
 agaacagatc aggaattcta aataaatttc ccagttaaag attattgtga cttaactgtta 1860  
 tataaacata ttttataact ttattgaaag gggacacctg tacattcttc catcatcact 1920  
 gttaaagacaa ataaatgatt atattcacaa aaaaaaaaaa aaaaaaa 1966

*JKB*

<210> 6  
<211> 414  
<212> PRT  
<213> Unknown

<400> 6

Met	Arg	Ala	His	Pro	Gly	Gly	Gly	Arg	Cys	Cys	Pro	Glu	Gln	Glu	Glu
1				5					10			15			
Gly	Glu	Ser	Ala	Ala	Gly	Gly	Ser	Gly	Ala	Gly	Gly	Asp	Ser	Ala	Ile
					20				25					30	
Glu	Gln	Gly	Gly	Gln	Gly	Ser	Ala	Leu	Ala	Pro	Ser	Pro	Val	Ser	Gly
						35			40				45		
Val	Arg	Arg	Glu	Gly	Ala	Arg	Gly	Gly	Arg	Gly	Arg	Gly	Arg	Trp	
						50			55			60			
Lys	Gln	Ala	Ala	Arg	Gly	Gly	Val	Cys	Gly	Arg	Gly	Arg	Gly	Arg	
						65			70			75		80	
Gly	Arg	Gly													
						85			90			95			
Pro	Gln	Ser	Gly	Gly	Ser	Gly	Leu	Gly	Gly	Asp	Gly	Gly	Gly	Ala	
						100			105			110			
Gly	Gly	Cys	Gly	Val	Gly	Ser	Gly	Gly	Gly	Val	Ala	Pro	Arg	Arg	Asp
						115			120			125			
Pro	Val	Pro	Phe	Pro	Ser	Gly	Ser	Ser	Gly	Pro	Gly	Pro	Arg	Gly	Pro
						130			135			140			
Arg	Ala	Thr	Glu	Ser	Gly	Lys	Arg	Met	Asp	Cys	Pro	Ala	Leu	Pro	Pro
						145			150			155		160	
Gly	Trp	Lys	Lys	Glu	Glu	Val	Ile	Arg	Lys	Ser	Gly	Leu	Ser	Ala	Gly
						165			170			175			
Lys	Ser	Asp	Val	Tyr	Tyr	Phe	Ser	Pro	Ser	Gly	Lys	Lys	Phe	Arg	Ser
						180			185			190			
Lys	Pro	Gln	Leu	Ala	Arg	Tyr	Leu	Gly	Asn	Ala	Val	Asp	Leu	Ser	Ser
						195			200			205			
Phe	Asp	Phe	Arg	Thr	Gly	Lys	Met	Met	Pro	Ser	Lys	Leu	Gln	Lys	Asn
						210			215			220			
Lys	Gln	Arg	Leu	Arg	Asn	Asp	Pro	Leu	Asn	Gln	Asn	Lys	Gly	Lys	Pro
						225			230			235		240	
Asp	Leu	Asn	Thr	Thr	Leu	Pro	Ile	Arg	Gln	Thr	Ala	Ser	Ile	Phe	Lys
						245			250			255			
Gln	Pro	Val	Thr	Lys	Phe	Thr	Asn	His	Pro	Ser	Asn	Lys	Val	Lys	Ser
						260			265			270			
Asp	Pro	Gln	Arg	Met	Asn	Glu	Gln	Pro	Arg	Gln	Leu	Phe	Trp	Glu	Lys
						275			280			285			
Arg	Leu	Gln	Gly	Leu	Ser	Ala	Ser	Asp	Val	Thr	Glu	Gln	Ile	Ile	Lys
						290			295			300			

Thr Met Glu Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Gly Ser Asn  
 305 310 315 320  
 Asp Glu Thr Leu Leu Ser Ala Val Ala Ser Ala Leu His Thr Ser Ser  
 325 330 335  
 Ala Pro Ile Thr Gly Gln Val Ser Ala Ala Val Glu Lys Asn Pro Ala  
 340 345 350  
 Val Trp Ieu Asn Thr Ser Gln Pro Leu Cys Lys Ala Phe Ile Val Thr  
 355 360 365  
 Asp Glu Asp Ile Arg Lys Gln Glu Glu Arg Val Gln Gln Val Arg Lys  
 370 375 380  
 Lys Leu Glu Glu Ala Leu Met Ala Asp Ile Leu Ser Arg Ala Ala Asp  
 385 390 395 400  
 Thr Glu Glu Val Asp Ile Asp Met Asp Ser Gly Asp Glu Ala  
 405 410

<210> 7  
 <211> 2392  
 <212> DNA  
 <213> Unknown

<400> 7

agcggggccga ggagccgggc gcaatggagc ggaagagggtg ggagtgcccg gcgctccgc 60  
 agggctggga gagggaaagaa gtgcccagaa ggtcggggct gtcggccggc cacaggatg 120  
 tctttacta tagcccgagc gggaaagaat tccgcagcaa gcccgcagctg gcgcgcattacc 180  
 tggcggttc catggacactg agacaccccg acttccgcac gggcaagatg ctgatgagca 240  
 agatgaacaa gagccgcag cgcgtgcgt acgactcc tc caaccagggtc aaggccaagc 300  
 ccgacccgtaa cacggcgctg cccgtgcgc agacggcgta catcttcaag cagccggta 360  
 ccaagattac caaccacccc agcaacaagg tcaagagcga cccgcagaag gcggtgacc 420  
 agccgcgcca gctttctgg gagaagaagg tgagccgcctt gaacgcctt gacattgt 480  
 aggagcttgtt caagaccatg gaccccttca agggctgcg ggggtggga cctggctgca 540  
 cggatgagac gctgtgtcg gccatcgca gcccgcgtca cactagcacc atgccccatca 600  
 cggacacgtt ctggccgcgt gtggagaagg accccggcgat atggctcaac accacgcagc 660  
 ccctgtgcaa agccttcatg gtgaccgcg aggacatcg gaagcaggaa gagctgggt 720  
 agcagggtcg gaaaggcgatc gaggaggcgc tgatggccg catgtggcg cacgtggagg 780  
 agctggcccg tgacggggag gcccgcgtgg acaaggcctg cgctgaggac gacgacgagg 840  
 aagacgagga ggaggaggag gaggagcccg accccggacc ggagatggag cacgtctagg 900  
 gcagaggccc tgccgagagc ccgtgtgcg tgatggccg gcctgcagac gcggccctcg 960  
 gccccacgtg aaccaggctc gggccgcgaa cccagcctt gagacaccca ggaggaaggc 1020  
 cgtgtctcg gctcccttcc cggccgcgtcc ccactcccg gggctcgcc gcacacagct 1080  
 gggctgccc ccacccaaaa gaccctccac gtcgtccctc tacagatgtcc ggctcgcc 1140  
 agtgcgggt gtcctggc cctgcctggc tccctaaatcgat ctttgggtcc gaggccagct 1200  
 cctcccatg cccgtgtcc cagtccttgc agactggaga gcaaggccagca ggtgcccggc 1260  
 agtcggcgc cacggcttc tgacagctgg gagggtttcc cggtctggag gcgtatgttt 1320  
 gaaactcaca tcacccactg tgccgcgtga ggacggact ctggctgtct gtggggggca 1380  
 tgcaggacgg cgccactctc tgccctgcctt tgccgcgtgg ggtgcacac agccctaccg 1440  
 tgcctgatgt gctgtcccgat ggaggccgccttcc ctccttcactt aatgtaaaca cagtcgaggc 1500  
 acgtcatcg gcaaggcccttcc ctgtgtgccttcc acggccagcttcc tggcttcgtaa aaaccaact 1560  
 ccagccgcgtg ccaggccggat cttggctgcgc cggcgctgcgca agaatgtcc actgcccagcc 1620  
 ggccccccctg cctcggtttcc ctttctgttt agtggcgaca caggcaccca gctttgggt 1680  
 ggtgcgtacg ctcccaaggatggccgatc cactgggaca ggtgcgtggcttcccagacgtct 1740  
 cctcgaggtg cccagcttc caggccgcgtt ctggcccaag gctgttttgc gggatctgt 1800  
 cctaaccatc ccaggccctt ggcgcaggccgatc ggttccaaagc cacagacgcgatc tgcccccgt 1860  
 ggactttgcg gcaaggccctt ggggtgccttc ctggcccttgc cttggcccttgc gagggttcc 1920  
 aacgggtggg ttcawtggcc tggcccvagc gagcccccac ctgcatttgc ctttagggcc 1980  
 tagagagggc ctgtcccgatc gtcgcggccatc ccaaggatct ggtgcgtgc ccaggggggac 2040  
 tgcgtggccaa gatgcgtccccc tgcgtggccatc ctgtgcaccat ccctgtatggg gcctgaccgc 2100  
 gggagctgag gaagcgcgc tccaccgtcttcccttcaag gacccgcata gaggcagtgg 2160

gctggcagct tcctgctgct ccctgtcaga gtcaaagcac aaatccttag gacgggctca 2220  
 agggccaggg cagccgaggg aagctccagg tggggaccac gtcttcttgaa gtttgtgcc 2280  
 cactggctgg gacccttgc agtgggtgg cctccctct gtctgcctgg tggagggagc 2340  
 cgtggcgta gggacgtgac tgaataaagc caccatgggt ggatgtgctt gg 2392

<210> 8  
 <211> 285  
 <212> PRT  
 <213> Unknown

<400> 8

Met Glu Arg Lys Arg Trp Glu Cys Pro Ala Leu Pro Gln Gly Trp Glu  
 1 5 10 15  
 Arg Glu Glu Val Pro Arg Arg Ser Gly Leu Ser Ala Gly His Arg Asp  
 20 25 30  
 Val Phe Tyr Tyr Ser Pro Ser Gly Lys Lys Phe Arg Ser Lys Pro Gln  
 35 40 45  
 Leu Ala Arg Tyr Leu Gly Gly Ser Met Asp Leu Ser Thr Phe Asp Phe  
 50 55 60  
 Arg Thr Gly Lys Met Leu Met Asn Lys Met Asn Lys Ser Arg Gln Arg  
 65 70 75 80  
 Val Arg Tyr Asp Ser Ser Asn Gln Val Lys Gly Lys Pro Asp Leu Asn  
 85 90 95  
 Thr Ala Leu Pro Val Arg Gln Thr Ala Ser Ile Phe Lys Gln Pro Val  
 100 105 110  
 Thr Lys Ile Thr Asn His Pro Ser Asn Lys Val Lys Ser Asp Pro Gln  
 115 120 125  
 Lys Ala Val Asp Gln Pro Arg Gln Leu Phe Trp Glu Lys Lys Leu Ser  
 130 135 140  
 Gly Leu Ser Ala Phe Asp Ile Ala Glu Glu Leu Val Arg Thr Met Asp  
 145 150 155 160  
 Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Gly Cys Thr Asp Glu Thr  
 165 170 175  
 Leu Leu Ser Ala Ile Ala Ser Ala Leu His Thr Ser Thr Leu Pro Ile  
 180 185 190  
 Thr Gly Gln Leu Ser Ala Ala Val Glu Lys Asn Pro Gly Val Trp Leu  
 195 200 205  
 Asn Thr Ala Gln Pro Leu Cys Lys Ala Phe Met Val Thr Asp Asp Asp  
 210 215 220  
 Ile Arg Lys Gln Glu Glu Leu Val Gln Gln Val Arg Lys Arg Leu Glu  
 225 230 235 240  
 Glu Ala Leu Met Ala Asp Met Leu Ala His Val Glu Glu Leu Ala Arg  
 245 250 255  
 Asp Gly Glu Ala Pro Leu Asp Lys Ala Cys Ala Glu Glu Glu Glu  
 260 265 270  
 Glu Glu Glu Glu Glu Pro Glu Pro Glu Arg Val  
 275 280 285

<210> 9  
 <211> 17  
 <212> DNA  
 <213> Unknown

<400> 9

ctggcaagag cgatgtc

*Sukhi*

<210> 10  
<211> 22  
<212> DNA  
<213> Unknown

<400> 10  
agtctggttt acccttattt tg

22